LH

PCT09

RAW SEQUENCE LISTING   DATE: 02/28/2002
PATENT APPLICATION: US/09/831,180A   TIME: 08:53:34

Input Set : N:\Crf3\02132002\I831180A.raw
Output Set: N:\CRF3\02282002\I831180A.raw

```
 1 <110> APPLICANT: Chiaki Senoo
 2        Mariko Numata
 3 <120> TITLE OF INVENTION: Novel Trypsin Family Serine Proteases
 4 <130> FILE REFERENCE: 50026/027001
 5 <140> CURRENT APPLICATION NUMBER: US/09/831,180A
 6 <141> CURRENT FILING DATE: 2001-05-03
 7 <150> PRIOR APPLICATION NUMBER: PCT/JP99/06111
 8 <151> PRIOR FILING DATE: 1999-11-02
 9 <150> PRIOR APPLICATION NUMBER: JP 1998-313366
10 <151> PRIOR FILING DATE: 1998-11-04
11 <160> NUMBER OF SEQ ID NOS: 53
12 <170> SOFTWARE: FastSEQ for Windows Version 4.0
14 <210> SEQ ID NO: 1
15 <211> LENGTH: 1033
16 <212> TYPE: DNA
17 <213> ORGANISM: Mus musculus
18 <220> FEATURE:
19 <221> NAME/KEY: CDS
20 <222> LOCATION: (48)...(1010)
21 <400> SEQUENCE: 1
22        cctgcctcag tgttggagct ccccattgct gatgtgcagg caagccg atg aaa cga      56
23                                                              Met Lys Arg
24                                                              1
25        tgg aag gac aga aga aca ggc ctg ttg ctg cca ttg gtc ctc ctg ttg     104
26        Trp Lys Asp Arg Arg Thr Gly Leu Leu Leu Pro Leu Val Leu Leu Leu
27            5                  10                  15
28        ttt ggg gca tgt agc tca ctg gca tgg gta tgt ggc cgg cga atg agt     152
29        Phe Gly Ala Cys Ser Ser Leu Ala Trp Val Cys Gly Arg Arg Met Ser
30            20                  25                  30                  35
31        agc aga tcc caa caa ctt aac aat gct tct gct atc gtg gaa ggc aaa     200
32        Ser Arg Ser Gln Gln Leu Asn Asn Ala Ser Ala Ile Val Glu Gly Lys
33                        40                  45                  50
34        cct gct tct gct atc gtg gga ggc aaa cct gca aac atc ttg gag ttc     248
35        Pro Ala Ser Ala Ile Val Gly Gly Lys Pro Ala Asn Ile Leu Glu Phe
36                        55                  60                  65
37        ccc tgg cat gtg ggg att atg aat cat ggt agt cat ctc tgt ggg gga     296
38        Pro Trp His Val Gly Ile Met Asn His Gly Ser His Leu Cys Gly Gly
39                        70                  75                  80
40        tct att ctc aat gag tgg tgg gtt cta tct gca tcc cat tgc ttc gac     344
41        Ser Ile Leu Asn Glu Trp Trp Val Leu Ser Ala Ser His Cys Phe Asp
42                        85                  90                  95
43        caa cta aac aac tct aaa ttg gag atc att cat ggc act gaa gac ctc     392
44        Gln Leu Asn Asn Ser Lys Leu Glu Ile Ile His Gly Thr Glu Asp Leu
```

R-5

ENTERED

```
45          100                     105                     110                     115
46          agc aca aag ggc ata aag tat cag aaa gtg gac aag tta ttc ttg cac      440
47          Ser Thr Lys Gly Ile Lys Tyr Gln Lys Val Asp Lys Leu Phe Leu His
48                       120                     125                     130
49          cca aag ttt gat gac tgg ctc ctg gac aac gac ata gct ttg ctc ttg      488
50          Pro Lys Phe Asp Asp Trp Leu Leu Asp Asn Asp Ile Ala Leu Leu Leu
51                       135                     140                     145
52          ctc aaa tcc cca tta aac ttg agt gtc aac agg ata cct atc tgc act      536
53          Leu Lys Ser Pro Leu Asn Leu Ser Val Asn Arg Ile Pro Ile Cys Thr
54                   150                     155                     160
55          tca gaa atc tct gac ata cag gca tgg agg aac tgc tgg gtg aca gga      584
56          Ser Glu Ile Ser Asp Ile Gln Ala Trp Arg Asn Cys Trp Val Thr Gly
57               165                     170                     175
58          tgg ggc att act aat act agt gaa aaa gga gtc caa ccc aca att ctg      632
59          Trp Gly Ile Thr Asn Thr Ser Glu Lys Gly Val Gln Pro Thr Ile Leu
60          180                     185                     190                     195
61          cag gca gtc aaa gtg gat ctg tac aga tgg gat tgg tgt ggc tat att      680
62          Gln Ala Val Lys Val Asp Leu Tyr Arg Trp Asp Trp Cys Gly Tyr Ile
63                       200                     205                     210
64          ttg tct cta tta acc aag aat atg ctg tgt gct ggg act caa gat cct      728
65          Leu Ser Leu Leu Thr Lys Asn Met Leu Cys Ala Gly Thr Gln Asp Pro
66                   215                     220                     225
67          ggg aag gat gcc tgc cag ggc gac agt gga gga gct ctc gtt tgc aac      776
68          Gly Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Ala Leu Val Cys Asn
69                   230                     235                     240
70          aaa aag aga aac aca gcc att tgg tac cag gtg ggc att gtc agc tgg      824
71          Lys Lys Arg Asn Thr Ala Ile Trp Tyr Gln Val Gly Ile Val Ser Trp
72               245                     250                     255
73          ggc atg ggc tgt ggc aag aag aat ctg cca gga gta tac acc aag gtg      872
74          Gly Met Gly Cys Gly Lys Lys Asn Leu Pro Gly Val Tyr Thr Lys Val
75          260                     265                     270                     275
76          tca cac tat gtg agg tgg atc agc aag cag aca gcg aag gcg ggg agg      920
77          Ser His Tyr Val Arg Trp Ile Ser Lys Gln Thr Ala Lys Ala Gly Arg
78                       280                     285                     290
79          cct tat atg tat gag cag aac tct gcg tgc cct ttg gtg ctc tct tgc      968
80          Pro Tyr Met Tyr Glu Gln Asn Ser Ala Cys Pro Leu Val Leu Ser Cys
81                   295                     300                     305
82          cgg gct atc ttg ttc cta tat ttt gta atg ttt ctt cta acc              1010
83          Arg Ala Ile Leu Phe Leu Tyr Phe Val Met Phe Leu Leu Thr
84                   310                     315                     320
85          tgatgattaa acgtgagact gcc                                            1033
87  <210> SEQ ID NO: 2
88  <211> LENGTH: 321
89  <212> TYPE: PRT
90  <213> ORGANISM: Mus musculus
91  <400> SEQUENCE: 2
92          Met Lys Arg Trp Lys Asp Arg Arg Thr Gly Leu Leu Leu Pro Leu Val
93          1                   5                       10                      15
94          Leu Leu Leu Phe Gly Ala Cys Ser Ser Leu Ala Trp Val Cys Gly Arg
```

```
 95                    20                   25                   30
 96        Arg Met Ser Ser Arg Ser Gln Gln Leu Asn Asn Ala Ser Ala Ile Val
 97                35                   40                   45
 98        Glu Gly Lys Pro Ala Ser Ala Ile Val Gly Gly Lys Pro Ala Asn Ile
 99          50                   55                   60
100        Leu Glu Phe Pro Trp His Val Gly Ile Met Asn His Gly Ser His Leu
101        65                   70                   75                   80
102        Cys Gly Gly Ser Ile Leu Asn Glu Trp Trp Val Leu Ser Ala Ser His
103                85                   90                   95
104        Cys Phe Asp Gln Leu Asn Asn Ser Lys Leu Glu Ile Ile His Gly Thr
105                100                  105                  110
106        Glu Asp Leu Ser Thr Lys Gly Ile Lys Tyr Gln Lys Val Asp Lys Leu
107                115                  120                  125
108        Phe Leu His Pro Lys Phe Asp Asp Trp Leu Leu Asp Asn Asp Ile Ala
109                130                  135                  140
110        Leu Leu Leu Leu Lys Ser Pro Leu Asn Leu Ser Val Asn Arg Ile Pro
111        145                  150                  155                  160
112        Ile Cys Thr Ser Glu Ile Ser Asp Ile Gln Ala Trp Arg Asn Cys Trp
113                165                  170                  175
114        Val Thr Gly Trp Gly Ile Thr Asn Thr Ser Glu Lys Gly Val Gln Pro
115                180                  185                  190
116        Thr Ile Leu Gln Ala Val Lys Val Asp Leu Tyr Arg Trp Asp Trp Cys
117                195                  200                  205
118        Gly Tyr Ile Leu Ser Leu Leu Thr Lys Asn Met Leu Cys Ala Gly Thr
119                210                  215                  220
120        Gln Asp Pro Gly Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Ala Leu
121        225                  230                  235                  240
122        Val Cys Asn Lys Lys Arg Asn Thr Ala Ile Trp Tyr Gln Val Gly Ile
123                245                  250                  255
124        Val Ser Trp Gly Met Gly Cys Gly Lys Lys Asn Leu Pro Gly Val Tyr
125                260                  265                  270
126        Thr Lys Val Ser His Tyr Val Arg Trp Ile Ser Lys Gln Thr Ala Lys
127                275                  280                  285
128        Ala Gly Arg Pro Tyr Met Tyr Glu Gln Asn Ser Ala Cys Pro Leu Val
129          290                  295                  300
130        Leu Ser Cys Arg Ala Ile Leu Phe Leu Tyr Phe Val Met Phe Leu Leu
131        305                  310                  315                  320
132        Thr
```

134 <210> SEQ ID NO: 3
135 <211> LENGTH: 1034
136 <212> TYPE: DNA
137 <213> ORGANISM: Mus musculus
138 <220> FEATURE:
139 <221> NAME/KEY: CDS
140 <222> LOCATION: (69)...(1025)
141 <223> OTHER INFORMATION:
142 <221> NAME/KEY: misc_feature
143 <222> LOCATION: 10
144 <223> OTHER INFORMATION: n = A or C or G or T/U

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/09/831,180A

TIME: 08:53:34

Input Set : N:\Crf3\02132002\I831180A.raw
Output Set: N:\CRF3\02282002\I831180A.raw

```
    145  <400> SEQUENCE: 3
W-->  146    cccacgcgtn cggttgtatc aatgtgggca gggcatcaag gcaggcacca ctgcactgga  60
      147    atgacaac atg atg ctc cca ctt cta att gca ctg ctc atg gct tcc aag  110
      148             Met Met Leu Pro Leu Leu Ile Ala Leu Leu Met Ala Ser Lys
      149              1               5                  10
      150    gga caa gct aag gac cag caa gaa tca gtt ctg tgt ggc cac aga cct  158
      151    Gly Gln Ala Lys Asp Gln Gln Glu Ser Val Leu Cys Gly His Arg Pro
      152     15                  20                  25                  30
      153    gcc ttc cca aac tca tca tgg ctg cca ttg cgg gag ctg ctt gag gtc  206
      154    Ala Phe Pro Asn Ser Ser Trp Leu Pro Leu Arg Glu Leu Leu Glu Val
      155                         35                  40                  45
      156    cag cat ggt gag ttc cca tgg caa gtg agt atc cag atg ctt ggg aaa  254
      157    Gln His Gly Glu Phe Pro Trp Gln Val Ser Ile Gln Met Leu Gly Lys
      158                         50                  55                  60
      159    cac ctg tgt gga ggc tcc atc atc cac cgg tgg tgg gtt ctg aca gca  302
      160    His Leu Cys Gly Gly Ser Ile Ile His Arg Trp Trp Val Leu Thr Ala
      161                     65                  70                  75
      162    gca cac tgc ttc ccg aga acc cta tta gaa ctg gta gca gtc aat gtc  350
      163    Ala His Cys Phe Pro Arg Thr Leu Leu Glu Leu Val Ala Val Asn Val
      164                 80                  85                  90
      165    act gtg gtc atg gga atc aag act ttc agt gac acc aac tta gag aga  398
      166    Thr Val Val Met Gly Ile Lys Thr Phe Ser Asp Thr Asn Leu Glu Arg
      167     95                  100                 105                 110
      168    aaa caa gtg cag aag atc att gct cac aga gac tac aaa ccg ccc gac  446
      169    Lys Gln Val Gln Lys Ile Ile Ala His Arg Asp Tyr Lys Pro Pro Asp
      170                     115                 120                 125
      171    ctt gac agc gac ctc tgc ctg ctc cta ctt gcc acg cca atc caa ttc  494
      172    Leu Asp Ser Asp Leu Cys Leu Leu Leu Leu Ala Thr Pro Ile Gln Phe
      173                     130                 135                 140
      174    aat aaa gac aaa atg ccc atc tgc ctg cca cag agg gag aac tcc tgg  542
      175    Asn Lys Asp Lys Met Pro Ile Cys Leu Pro Gln Arg Glu Asn Ser Trp
      176                 145                 150                 155
      177    gac cgg tgc tgg atg tca gag tgg gca tat act cat ggc cat ggt tca  590
      178    Asp Arg Cys Trp Met Ser Glu Trp Ala Tyr Thr His Gly His Gly Ser
      179                 160                 165                 170
      180    gcc aaa ggc tca aac atg cac ctg aag aag ctc agg gtg gtt cag att  638
      181    Ala Lys Gly Ser Asn Met His Leu Lys Lys Leu Arg Val Val Gln Ile
      182     175                 180                 185                 190
      183    agc tgg agg aca tgt gcg aag agg gtg act cag ctc tcc agg aac atg  686
      184    Ser Trp Arg Thr Cys Ala Lys Arg Val Thr Gln Leu Ser Arg Asn Met
      185                     195                 200                 205
      186    ctt tgt gct tgg aag gaa gtg ggc acc aac ggc aag tgc cag gga gac  734
      187    Leu Cys Ala Trp Lys Glu Val Gly Thr Asn Gly Lys Cys Gln Gly Asp
      188                     210                 215                 220
      189    agc ggg gca ccc atg gtc tgt gct aac tgg gag act cgg aga ctc ttt  782
      190    Ser Gly Ala Pro Met Val Cys Ala Asn Trp Glu Thr Arg Arg Leu Phe
      191                     225                 230                 235
      192    caa gtg ggt gtc ttc agc tgg ggc ata act tca gga tcc agg ggg agg  830
      193    Gln Val Gly Val Phe Ser Trp Gly Ile Thr Ser Gly Ser Arg Gly Arg
```

RAW SEQUENCE LISTING   DATE: 02/28/2002
PATENT APPLICATION: US/09/831,180A   TIME: 08:53:34

Input Set : N:\Crf3\02132002\I831180A.raw
Output Set: N:\CRF3\02282002\I831180A.raw

```
194        240              245              250
195    cca ggc att ttt gtg tct gtg gct cag ttt atc cca tgg atc ctg gag    878
196    Pro Gly Ile Phe Val Ser Val Ala Gln Phe Ile Pro Trp Ile Leu Glu
197    255              260              265              270
198    gag aca caa agg gag gga cga gcc ctt gcc ctc tca aag gcc tca aaa    926
199    Glu Thr Gln Arg Glu Gly Arg Ala Leu Ala Leu Ser Lys Ala Ser Lys
200              275              280              285
201    agt ctc ttg gct ggc agt cca cgc tac cat ccc ata ttg cta agc atg    974
202    Ser Leu Leu Ala Gly Ser Pro Arg Tyr His Pro Ile Leu Leu Ser Met
203              290              295              300
204    ggc tct caa ata ctg ctt gct gcc ata ttt tct gat gat aaa tca aat    1022
205    Gly Ser Gln Ile Leu Leu Ala Ala Ile Phe Ser Asp Asp Lys Ser Asn
206              305              310              315
207    tgc taagctctg                                                      1034
208    Cys
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 319
212 <212> TYPE: PRT
213 <213> ORGANISM: Mus musculus
214 <400> SEQUENCE: 4
215    Met Met Leu Pro Leu Leu Ile Ala Leu Leu Met Ala Ser Lys Gly Gln
216    1               5               10              15
217    Ala Lys Asp Gln Gln Glu Ser Val Leu Cys Gly His Arg Pro Ala Phe
218              20              25              30
219    Pro Asn Ser Ser Trp Leu Pro Leu Arg Glu Leu Leu Glu Val Gln His
220              35              40              45
221    Gly Glu Phe Pro Trp Gln Val Ser Ile Gln Met Leu Gly Lys His Leu
222         50              55              60
223    Cys Gly Gly Ser Ile Ile His Arg Trp Trp Val Leu Thr Ala Ala His
224    65              70              75              80
225    Cys Phe Pro Arg Thr Leu Leu Glu Leu Val Ala Val Asn Val Thr Val
226              85              90              95
227    Val Met Gly Ile Lys Thr Phe Ser Asp Thr Asn Leu Glu Arg Lys Gln
228              100             105             110
229    Val Gln Lys Ile Ile Ala His Arg Asp Tyr Lys Pro Pro Asp Leu Asp
230              115             120             125
231    Ser Asp Leu Cys Leu Leu Leu Leu Ala Thr Pro Ile Gln Phe Asn Lys
232              130             135             140
233    Asp Lys Met Pro Ile Cys Leu Pro Gln Arg Glu Asn Ser Trp Asp Arg
234    145             150             155             160
235    Cys Trp Met Ser Glu Trp Ala Tyr Thr His Gly His Gly Ser Ala Lys
236              165             170             175
237    Gly Ser Asn Met His Leu Lys Lys Leu Arg Val Val Gln Ile Ser Trp
238              180             185             190
239    Arg Thr Cys Ala Lys Arg Val Thr Gln Leu Ser Arg Asn Met Leu Cys
240              195             200             205
241    Ala Trp Lys Glu Val Gly Thr Asn Gly Lys Cys Gln Gly Asp Ser Gly
242              210             215             220
243    Ala Pro Met Val Cys Ala Asn Trp Glu Thr Arg Arg Leu Phe Gln Val
```

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/831,180A

DATE: 02/28/2002
TIME: 08:53:35

Input Set : **N:\Crf3\02132002\I831180A.raw**
Output Set: **N:\CRF3\02282002\I831180A.raw**

L:5 M:270 C: Current Application Number differs, Wrong Format
L:146 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:1012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47
L:1039 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49